Received December 21, 2018, accepted January 8, 2019, date of publication January 21, 2019, date of current version February 6, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2893448*

# XGBoost-Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System

**MINGHUA CHEN[1], QUNYING LIU[1], (Member, IEEE), SHUHENG CHEN[2], (Member, IEEE), YICEN LIU[1], CHANG-HUA ZHANG[2], (Member, IEEE), AND RUIHUA LIU[3]**

[1]School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[3]Skills Training Center, State Grid Sichuan Electric Power Company, Chengdu 611133, China

Corresponding author: Qunying Liu (lqy1206@126.com)

This work was supported in part by the Natural Science Foundation of China under Grant NSFC51677020, in part by the China Postdoctoral Science Foundation under Grant 2015M572457, in part by the Provincial Key Laboratory of Power Electronics Energy Saving Technology and Equipment under Grant szjj2016-093, and in part by the 2018 Fundamental Research Funds for the Central Universities.

**ABSTRACT** The artificial intelligence (AI) techniques have been widely used in the transient stability analysis of a power system. They are recognized as the most promising approaches for predicting the post-fault transient stability status with the use of phasor measurement units data. However, the popular AI methods used for power systems are often "black boxes," which result in the poor interpretation of the model. In this paper, a transient stability prediction method based on extreme gradient boosting is proposed. In this model, a decision graph and feature importance scores are provided to discover the relationship between the features of the power system and transient stability. Meanwhile, the key features are selected according to the feature importance scores to remove redundant variables. The simulation results on the New England 39-bus system have demonstrated the superiority of the proposed model over the prior methods in the computation speed and prediction accuracy. Finally, an algorithm is proposed to interpret the prediction results for a specific fault of the power system, which further improves the interpretability of the model and makes it attractive for real-time transient stability prediction.

**INDEX TERMS** Feature importance scores, model interpretation, XGBoost model, transient stability prediction.

## I. INTRODUCTION

With the development of the power grid, the dynamic characteristics of the power system become more complex and the stability analysis and control of the power system is more difficult to be performed. When the power system suffers various disturbances, the stability problems will be caused to a different extent. Thus, it is of great significance to predict the transient stability of power system quickly and take emergency control actions after suffering a severe disturbance. The commonly used methods of transient stability assessment (TSA) include time domain simulation and direct methods [1]. The time domain simulation method is currently the most robust method available for transient stability assessment. However, it cannot make a trade-off between time-consuming and resource-intensive while using the parallel computation-based time domain simulation [2]. The direct methods include the transient energy function [3], the extended equal area criterion [4], etc. They work fast but the results are conservative. These methods cannot fully meet the actual requirements of online TSA [5].

In recent years, artificial intelligence (AI) has been introduced into power system, which has improved the solutions in many applications. As a competitive technology for accomplishing these applications, PMU provides the synchronized parameters measurements with high sampling rate in the milliseconds range. With the sampling data from PMU, AI can be used to create the mapping between features of power system and post-fault stability status.

With the development of machine learning theory, artificial neural network [6]–[8], support vector machines (SVM) [9]–[11], decision tree [12]–[14] and other classical models are widely applied to transient stability prediction.

Artificial neural network (ANN) based method is developed for quickly estimating the long-term voltage stability margin [6]. The grey wolf optimization and particle swarm optimization are introduced in [8] to train feedforward neural network. The work presented in [9] has shown that the transient stability status of power system suffering a large disturbance can be predicted timely by using the model of SVM method based on the measured state parameters of the generators. However, the number of input variables of the model is related to the scale of the power system. Much longer time will be spent to predict the transient stability when the scale of power system is large. The method of SVM with combinatorial trajectory inputs was trained to predict the transient stability status, and also give the credible area and incredible area of classifier [10]. In [12]–[14], an out-of-step detection technique based on decision tree has been proposed. The method presented in [12] has been applied to Iran national grid. In [15], the random forest and recursive feature elimination are proposed to select the key features of transient stability assessment. Based on [15], a weighted random forest is introduced to assess transient stability in [16]. In [17], SVMs are used as weak classifiers in adaptive boosting algorithm and they are further improved by a new weight updating strategy based on fuzzy clustering threshold technique. The ensemble learning machine is used to improve the prediction accuracy by combining a series of weak base learners into a strong one. However, with the expansion of power system, the amount of PMU data is increasing greatly. Thus, it is hard for the AI methods mentioned above to utilize the real-time data for online TSA because of the low prediction accuracy and long training time. Another reason is that many AI models are "black boxes", so that the decision rules obtained by the models cannot be understood from the perspective of human.

XGBoost is an efficient and scalable implementation of the Gradient Boosting Machine (GBM), which has been a competitive tool among artificial intelligence methods due to its features such as easy parallelism and high prediction accuracy [18]. Furthermore, the following advantages make it adaptable to deal with the transient stability prediction:

(1) In XGBoost model, multithreading parallel computing can be automatically called, which is faster than the traditional ensemble learning to predict the transient stability with large amounts of data in the actual power grid.

(2) That the regularization term addition to XGBoost, makes its generalization ability be improved, which makes up for the shortcoming that the decision tree is easy to be over-fitted.

(3) XGBoost is the tree structure model, which doesn't need to normalize the data collected by PMU in the power system. Furthermore, it can effectively deal with the missing values, which is suitable for PMU-based transient stability prediction to discover the relationship between features and transient stability.

This paper makes three main contributions to alleviate some of the drawbacks discussed above. First, the XGBoost model with high efficiency and accuracy in transient stability prediction is introduced. Second, feature selection strategies are implemented by two methods: correlation filtering and model-based feature selection. Third, the paper gives the explanation to the prediction result of the model in the perspective of human. This paper is organized as follows. Feature selection techniques are discussed in Section II. The principles of XGBoost are presented in Section III. Section IV and Section V respectively describe the model evaluation and the flow chart of transient stability prediction. Section VI presents some case studies and Section VII is the conclusion of this paper.

## II. FEATURE SELECTION PROCEDURE

PMU data is high-dimension time series, which is not suitable for large-scale system analysis. How to select useful information from massive data to generate raw input features is the premise of machine learning for transient stability prediction. The following principles should be followed when creating the raw input features: 1) features can reflect the transient stability well; 2) the number of features does not increase with the scale of the power system; 3) the calculation of features should be fast enough. According to principles, table 1 shows the 21 features selected with reference to experience [19]–[21] of the researchers in features selection. In table 1, $t_0$ is the beginning time of the fault, $t_c$ is the cutting time of the fault.

According to the features of table 1, the correlation between every two features is analyzed to determine whether the information among features is redundant. The correlation coefficient indicates the linear correlation between two random variable $f_i$ and $f_j$, and the formula is shown as:

$$c = \frac{\sum_{l=1}^{n}\left(f_i^{(l)} - \bar{f}_i\right)\left(f_j^{(l)} - \bar{f}_j\right)}{\sqrt{\sum_{l=1}^{n}\left(f_i^{(l)} - \bar{f}_i\right)^2 \sum_{l=1}^{n}\left(f_j^{(l)} - \bar{f}_j\right)^2}} \tag{1}$$

The values of $c$ range from $-1$ to 1. Usually, if $|c| > 0.5$, it indicates that the two features have a strong correlation; if $|c|$ is close to 0, it indicates that there is no linear correlation between the features. Correlation matrix diagrams between features are shown in Figure 1 and the values in the figure are obtained according to formula (1).

The correlation matrix is computed to check the linear relationship between the variables, which is used to identify the highly correlated variables. The darker the color in the graph is, the higher the correlation between the two features is. Two variables have extremely high correlations magnitudes which indicate that they are containing similar information. The correlation filtering is intended to remove the redundant variables. Some algorithms will be used to build unstable models and decelerate the training process if the model has redundant variables. Therefore, if a pair of variables has high correlations ($|c| > 0.98$), one of the two will be removed. Here, $f_0, f_1$ and $f_{19}$ are removed accordingly.

**TABLE 1. Feature sets.**

| Features | Description |
|---|---|
| $f_0$ | Maximum value of rotors acceleration at $t_0$ |
| $f_1$ | Maximum value of rotors kinetic energy at $t_c$ |
| $f_2$ | Rotor angle of generator that has largest acceleration at $t_0$ |
| $f_3$ | Rotor angle of generator that has largest kinetic energy at $t_c$ |
| $f_4$ | Kinetic energy of generator that has largest rotor angle at $t_c$ |
| $f_5$ | Total system 'energy adjustment' |
| $f_6$ | Minimum generator relative acceleration at $t_0$ |
| $f_7$ | Root mean square error of rotors acceleration at $t_0$ |
| $f_8$ | Average value of all rotor kinetic energy at $t_c$ |
| $f_9$ | Difference of max. and min. generator rotor angle at $t_c$ |
| $f_{10}$ | Difference of max. and min. kinetic energy at $t_c$ |
| $f_{11}$ | Difference of max. and min. generator angular velocity at $t_c$ |
| $f_{12}$ | Average value of all rotor acceleration at $t_0$ |
| $f_{13}$ | Sum of generator rotor mechanical power at $t_0$ |
| $f_{14}$ | Maximum active power impact of generator at $t_0$ |
| $f_{15}$ | Minimum active power impact of generator at $t_0$ |
| $f_{16}$ | Maximum reactive power impact of generator at $t_0$ |
| $f_{17}$ | Minimum reactive power impact of generator at $t_0$ |
| $f_{18}$ | Sum of the absolute values generator angular velocity |
| $f_{19}$ | Maximum generator angular velocity |
| $f_{20}$ | Impact level of fault to the system |

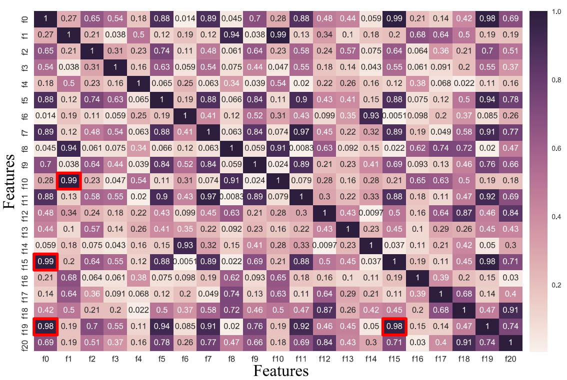



**FIGURE 1. Correlation matrix between features.**

## III. PRINCIPLES OF XGBOOST MODEL

XGBoost has been widely used in many fields to achieve state-of-the-art results on some data challenges (e.g., Kaggle competitions), which is a high effective scalable machine learning system for tree boosting [22]–[24]. XGBoost is optimized under the Gradient Boosting framework and developed by Chen and Guestrin [18], which is designed to be highly efficient, flexible and portable. The main idea of boosting is to combine a series of weak classifiers with low accuracy to build a strong classifier with better classification performance. If the weak learner for each step is based on the gradient direction of the loss function, it can be called the Gradient Boosting Machines.

Assuming that a data set is $D = \{(x_i, y_i) : i=1 \ldots n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R}\}$, we have $n$ samples with $m$ features. Let $\hat{y}_i$ be defined as the predict value by the model:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in F \tag{2}$$

where $f_k$ represents an independent regression tree and $f_k(x_i)$ denotes the prediction score given by the $k$-th tree to the $i$-th sample. The set of functions $f_k$ in the regression tree model can be learned by minimizing the objective function:

$$Obj = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega(f_k) \tag{3}$$

The $l$ herein is a training loss function, which measures the difference between the prediction $\hat{y}$ and the object $y_i$. To avoid over-fitting, the term $\Omega$ penalizes the complexity of the model:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{4}$$

where $\gamma$ and $\lambda$ are the degrees of regularization. $T$ and $w$ are the numbers of leaves and the scores on each leaf respectively.

The tree ensemble model can be trained in an additive manner. Let $\hat{y}_i^{(t)}$ be the prediction of the $i$-th instance at the $t$-th iteration, it needs to add $f_t$ to minimize the following objective:

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{5}$$

The equation (6) is obtained by using the second order Taylor expansion to simplify the equation (5) and remove the constant term:

$$Obj^{(t)} = \sum_{i=1}^{n} \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t) \tag{6}$$

where $g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ and $h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ are the first and the second order gradient on $l$. Then the objective is rewritten as:

$$\begin{aligned} Obj^{(t)} &= \sum_{i=1}^{n} \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \\ &= \sum_{j=1}^{T} \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \end{aligned} \tag{7}$$

where $I_j = \{i \mid q(x_i) = j\}$ denotes the instance set of leaf $j$. For a fixed tree structure $q$, the optimal weight $w_j^*$ of leaf $j$ and the corresponding optimal value can be calculated by:

$$w_j^* = -\frac{G_j}{H_j + \lambda} \tag{8}$$

$$Obj^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \lambda T \tag{9}$$

**TABLE 2. Confusion matrix.**

| Observed | Predicted | |
|---|---|---|
| | Stability | Instability |
| Stability | TP | FN |
| Instability | FP | TN |

where $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$, $obj$ presents the quality of a tree structure $q$. The smaller the value is, the better the structure of the tree is. Since it is impossible to enumerate all the tree structures, a greedy algorithm is used to add branches to the tree iteratively. $I_L$ and $I_R$ are the instance sets of the left and right nodes after split. By enumerating the feasible segmentation points and selecting the minimum target function and the maximum gain partition, the gain formula is shown as follows:

$$G = \frac{1}{2}\left[\frac{\left(\sum_{i \in I_L} g_i\right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_R} g_i\right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i\right)^2}{\sum_{i \in I} h_i + \lambda}\right] - \gamma \tag{10}$$

This formula is usually used in practice for evaluating the split candidates. The XGBoost model produces many simple trees, which are used to score a leaf node during splitting. The first, second and third term of the equation stand for the score on the left, right and the original leaf respectively. Moreover, the term$\gamma$ is the regularization on the additional leaf. It will be used in the training process.

## IV. MODEL OF TRANSIENT STABILITY ANALYSIS BASED ON XGBOOST

### A. CROSS VALIDATION

The simulation database is randomly divided into training set (80%) and test set (20%). K-fold cross-validation is used for the training set to find the best parameters of the model. 10-fold cross validation is commonly used in practice. In the 10-fold cross validation, the training set is split into ten parts of approximately equal size, in which nine parts are used for training and one part is used for validation. This process is repeated ten times iteratively and the average of these accuracy is taken as the expected prediction accuracy.

### B. CONSTRUCTION OF EFFECTIVENESS EVALUATION INDEX FOR TRANSIENT STABILITY PREDICTION

For a transient stability prediction model, the cores of the model are high accuracy rate and superior computation speed. It is hoped that the model can make a fast and accurate determination of the transient stability status of post-fault power system. The evaluation index proposed in this paper is based on these two aspects. To evaluate the accuracy of the classification model, the confusion matrix is listed in table 2. The intersection of the rows and columns show one of the four outcomes. For example, if we predict a case is stable, but it actually is instable, this is a false positive (FP).

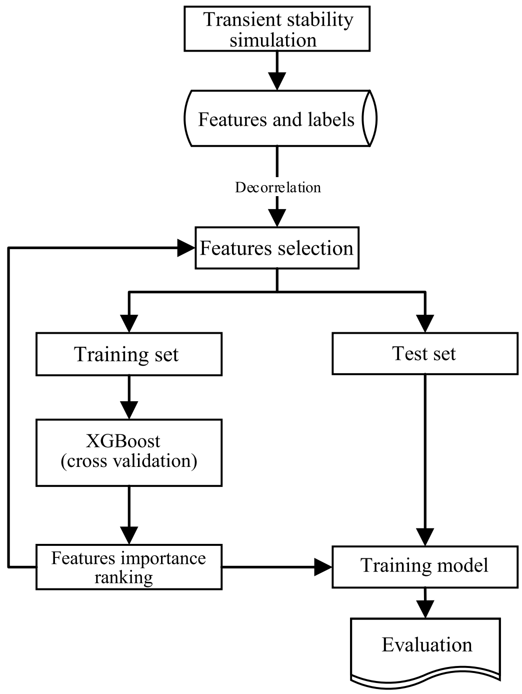



**FIGURE 2. XGBoost-based transient stability prediction process.**

It is considered that stable samples are less costly to predict as unstable, and the unstable samples are more costly to predict as stable. Once the instability case is omitted and predicted to be stable without taking any measures, it would perhaps lead to the disastrous consequences. Therefore, different evaluation indexes are considered as follows.

The missing alarm rate (MAR), which is used to measure the fraction of unstable samples, are predicted to be security. Here, the missing alarm rate is used to measure risk in transient stability status prediction of power system.

$$MAR = \frac{TP}{FP + TN} \tag{11}$$

The false alarm rate (FAR) measures the fraction of the forecasted insecurity events that doesn't occur. Its expression is shown as:

$$FAR = \frac{FN}{TP + FN} \tag{12}$$

However, the overall accuracy that measures the total classification accuracy is needed, it is expressed as:

$$ACC = \frac{TP + TN}{TP + FN + FP + TN} \tag{13}$$

## V. TRANSIENT STABILITY PREDICTION PROCESS

In this paper, an XGBoost-based transient stability prediction model is proposed. The flowchart of the proposal is depicted in Figure 2. The proposed model is implemented via the following steps: 1) generator operating parameters are obtained by simulating the expected faults, and all the features are calculated according to table 1; 2) correlation

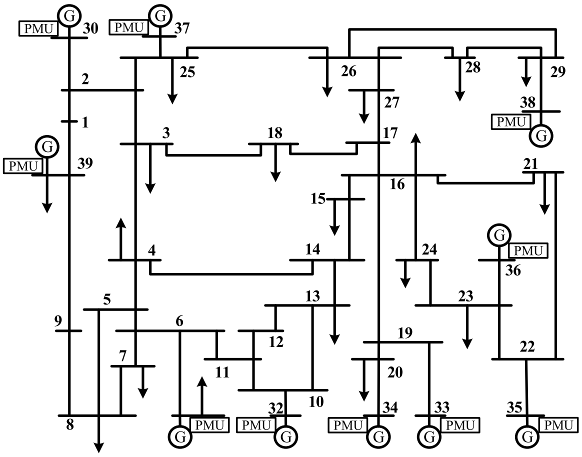



**FIGURE 3.** New England 10-machine 39-bus power system.

analysis of original feature set; 3) XGBoost prediction model is generated by training, and the trained transient stability prediction model is evaluated by the test set.

### 1) DATA PRE-PROCESSING

The time-domain datum including generators speeds, rotor angles and power are generated by the simulation of Power System Analysis Software Package (PSASP). All the raw input features are calculated according to table 1.

### 2) FEATURE SELECTION

Feature selection strategies are implemented by two methods: correlation filtering and the model-based feature selection. Correlation filtering is intended to remove the redundant features with extremely high correlations. The model-based feature selection calculates the feature importance scores of the model and filters redundant features with low scores because they are regarded as unimportant variables.

### 3) MODEL TRAINING AND EVALUATION

The initial step of the training process includes a grid search for finding the optimal parameters. The performance of the proposed transient stability prediction model is evaluated by the test set.

## VI. CASE STUDY

In this paper, the New England 10-machine 39-bus power system is used to illustrate the application of the proposed XGBoost model. First of all, the datum are collected by the simulation of PSASP. The contingencies considered are mainly three-phase to ground faults happened at 60 different locations. The above contingencies were repeated at five different kinds of loading levels (80%, 90%, 100%, 110%, 120%). At each loading level, five different kinds of active power of generators are changed randomly in the range of 80%-120%. The simulations assume that the start time of the fault is at 0.0s and the fault is cut off at 0.2s. The data sampling period is 0.01s, and 1200 samples are collected.

**TABLE 3.** Performance of different classification models.

| Models | Test set | | | Training time |
|---|---|---|---|---|
| | MAR | FAR | ACC | |
| XGBoost | 1.80% | 2.56% | 97.82% | 11 ms |
| RF | 3.04% | 4.68% | 96.14% | 25 ms |
| DT | 4.36% | 5.08% | 95.28% | 10 ms |
| SVM | 3.08% | 3.82% | 96.55% | 49 ms |
| BPNN | 4.64% | 6.96% | 94.20% | 122 ms |

A large number of PMUs that can monitor generator operating parameters are being developed for predicting transient stability in power system. The data collected will include rotor angle, angular velocity, active power, reactive power and mechanical power of generators. The stability of the system is judged by whether the relative rotor angle of any two generators is less than 360 degrees.

### A. CLASSIFIER PERFORMANCE COMPARISON

In order to evaluate the effectiveness of the proposal, a test between the XGBoost model and other popular machine learning models is further carried out. The performance of XGBoost model is compared with the other models, such as random forest (RF), decision tree (DT), support vector machines (SVM) and back propagation neural network (BPNN). The performance of different classification models is summarized in table 3.

According to table 3, it can be found that in these models, the XGBoost method has performed significantly better than other models by the indexes MAR, FAR and ACC. Since the decision tree is a single tree model, it takes a shorter training time but performs worse than XGBoost and RF.

### B. TRANSIENT PREDICTION MODEL ANALYSIS

In addition to training time and accuracy, interpretability is also a vital factor to the transient stability prediction model. However, interpretability tends to be neglected in many studies for the popular machine learning models (e.g., support vector machines and Neural network) which are inherent the ''black-box'' systems. As the tree boosting model, the interpretability of the XGBoost-based transient stability status prediction model mainly has reflected in two aspects: decision rules and feature importance score.

This XGBoost-based model generates 100 basic tree models, as depicted in Figure 4. The feature importance scores (F-score) denotes the number of times that a feature is used for splitting in the training process. The features are sorted in descending order of their relative importance scores, which is shown in Figure 5.

A higher score indicates that the corresponding feature is more important. Therefore, $f_9$ (difference of max. and min. generator rotor angle at $t_c$) is valuable and should be highlighted. By comparison, $f_{10}$ (difference of max. and min. kinetic energy at $t_c$) is not a necessary feature in the

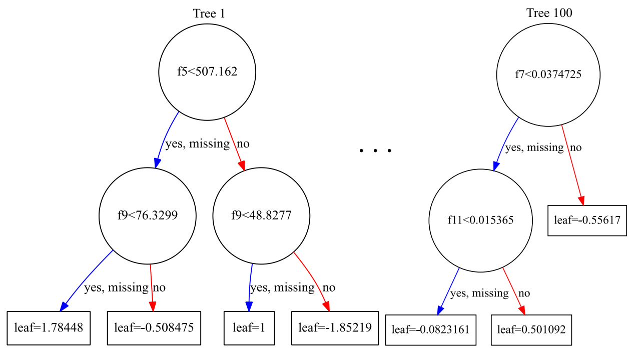



**FIGURE 4.** Decision chart of XGBoost-based transient stability prediction model.

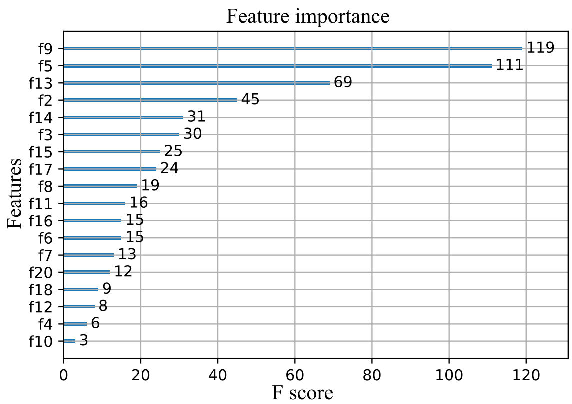



**FIGURE 5.** Diagram of feature importance scores.

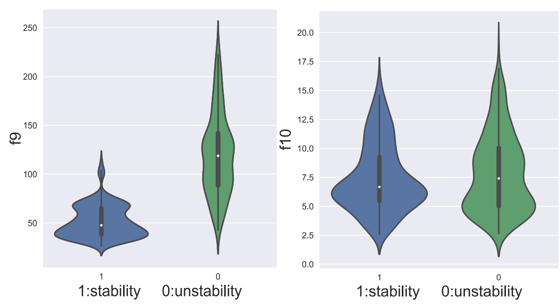



**FIGURE 6.** Distribution of $f_9$ and $f_{10}$ in stable and unstable conditions.

model and may be abandoned to speed up the training process of the model. Figure 6 shows the distribution of $f_9$ and $f_{10}$ in the stable and the unstable conditions. If one feature has the similar distribution under stable and unstable conditions, it means that this feature does not have the ability to judge transient stability. It is obvious that the distribution of $f_9$ is more favorable for transient stability classification.

According to the feature importance scores, three features with the highest scores are selected to draw the three-dimensional graph of the distribution of samples, as shown in Figure 7. In Figure 7, red dots indicate unstable samples and green dots indicate stable samples. It can be seen from

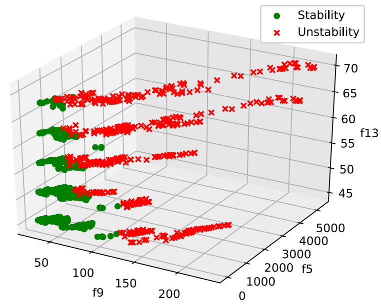



**FIGURE 7.** Three-dimensional distribution graph of samples under the features of $f_5$, $f_9$ and $f_{13}$.

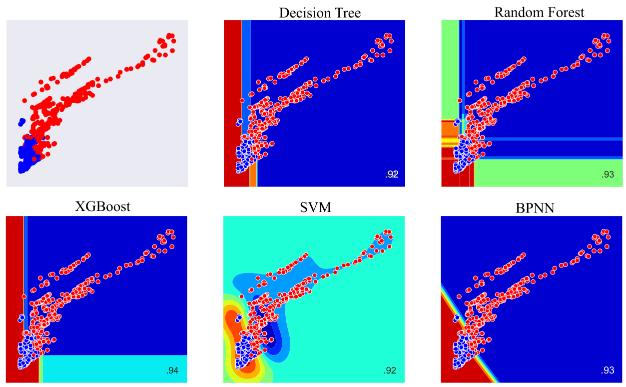



**FIGURE 8.** Decision boundary of different classification models under the two-dimensional graph of samples under features of $f_5$ and $f_9$. The classification accuracy is shown in the lower right corner of every figure.

the Figure 7 that the three features can distinctly distinguish unstable samples from stable samples. It means that this algorithm can identify the key features of the transient stability of power system.

For simplicity, two features ($f_5$ and $f_9$) are chosen to draw the decision boundaries based on the different classification models as shown in Figure 8. In Figure 8, the red dots represent the unstable samples, and the blue dots represent stable samples.

As is seen, the tree-based models have the similar decision boundaries. The XGBoost model has the highest accuracy among the models. It's easy to find out the decision boundary from the graph because it has only two features. If there are many features, it is difficult to find out the decision boundary and it is impossible to know why the model determines one case to be stable or unstable in power system. Hence, there appears a question that why we should trust the model? In other words, if we do not trust a classifier, we won't use it. Of course, the best way of earning the trust of a human would be that the AI can explain how it comes to a given decision. One approach is LIME, which stands for Local Interpretable Model-agnostic Explanations and is a tool that helps to understand and explain the decisions made by machine learning models [25]. This algorithm is used to explain the prediction results in the last part of this section.

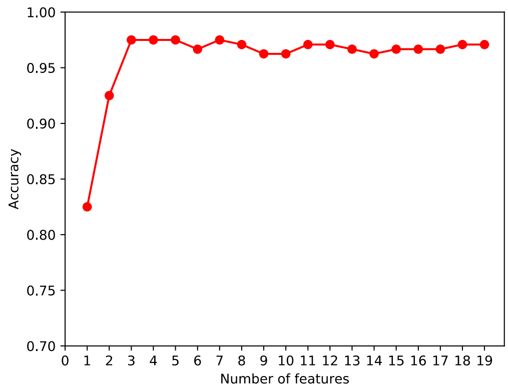



**FIGURE 9.** Relationship between numbers of features and predictive performance.

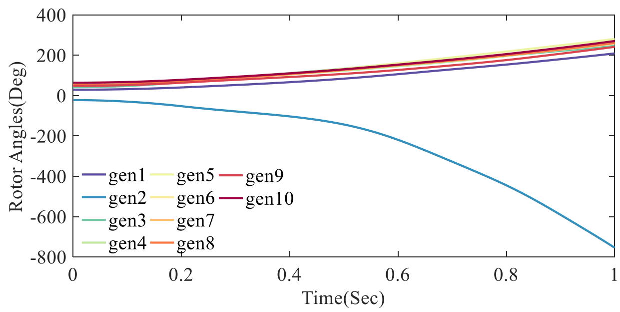



**FIGURE 10.** The rotor angle for Generator1-10 when fault occurred at line 5-7.

## C. PREDICTION MODEL PERFORMANCE WITH DIFFERENT NUMBER OF FEATURES

According to the ranking of feature importance scores, the features are added one by one from high to low and the corresponding prediction accuracy is calculated. The number of features and the prediction performance curve are shown in Figure 9. The prediction performance tends to be stable when the number of features reaches three.

## D. TRANSIENT STABILITY ANALYSIS OF A THREE-PHASE FAULT OCCURRED ON TRANSMISSION LINE

Assuming that the faults are occurred at line 5-7 and line 1-9 and they will last 0.2s respectively, the data of rotor angle, angular velocity, active power, reactive power and mechanical power of generators are collected. The curves of the generator angle are shown in Figure10 and Figure 11.

Traditionally, the most direct method to judge the stability of power system is to observe the rotor angle of generator. Therefore, Figure10 and Figure 11 show the change of rotor angle for generator 1-10 when the three-phase short-circuit fault happens respectively at line 1-9 and line 5-7. From Figure 10, it can be seen that the line 1-9 remains stable after the fault. While the fault occurs at line 5-7, the Gen2 reaches 360 degrees at about 0.73s, and loses its step. It means that the power system is unstable in this case. By this method based on XGBoost in this paper, the prediction results about the stability of power system after fault is shown in table 4.

In table 4, the samples used in XGBoost model come from the PMU. During this period that the three-phase short-circuit fault occurs and disappears, the PMU in plant collects all the operating parameters of generators. Based on these data, features are calculated. Under the operating condition that load level of the whole system reaches 120% and the fault at line 5-7 is cleared at 0.2s, the prediction results show that the system is unstable. However, with the operating condition that 90% load level and the fault at line 1-9 being cleared at 0.2s, the system is judged to be still stable by the prediction results. By comparing with the result from Figure10 and Figure 11, the same result has been obtained. However, the direct judgment by observing the rotor angle needs much time. Only when the rotor angle reaches some value, one can draw conclusions that if it is stable or not. In this case, it takes 0.73s to get the result. Furthermore, in Figure 10 or 11, it is not easy to judge if there is the abnormal generator in a short time if the monitor has no much experience. Hence, more time will be taken to observe the curve for a still stable system after fault. The results for the proposed method based on XGBoost show that only 2.81ms has been taken to get the conclusion when system is unstable after fault. However, less time 2.25ms has been taken to give the result when system is still stable after fault. Therefore, the XGBoost-based model shows an overwhelmed advantage

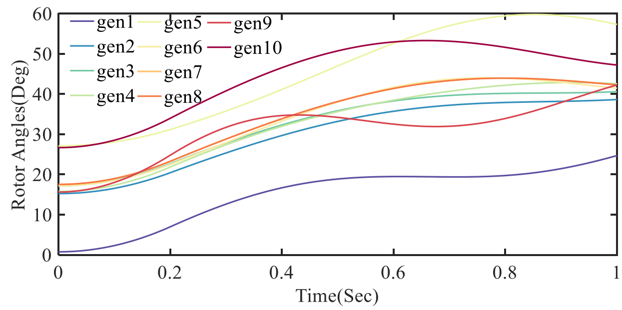



**FIGURE 11.** The rotor angle for Generator 1-10 when fault occurred at line 1-9.

**TABLE 4.** Prediction results of stability after fault.

| Location | Line 5-7 | Line 1-9 |
|---|---|---|
| Operation state | 120% load level, three-phase fault, cutting off fault at 0.2s | 90% load level, three-phase fault, cutting off fault at 0.2s |
| Features | $f_9$:130.7079 $f_5$:812.9402 $f_{13}$:45.0486… | $f_9$:26.9779 $f_5$:0.1929 $f_{13}$:45.5325… |
| Actual state | after 0.73s lose stability | stability |
| Predictive state | instability | stability |
| Time-consuming | 2.81 ms | 2.25 ms |

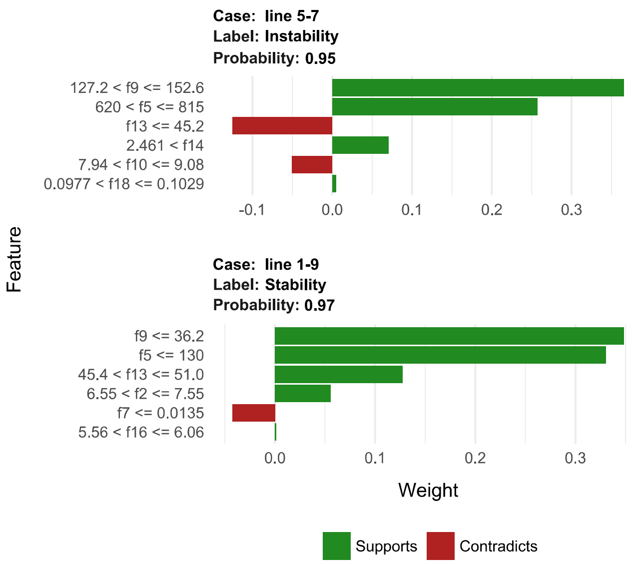



**FIGURE 12.** **LIME is used to explain the prediction results of the two cases. The header of each figure gives the case detail, in which class label is predicted and the corresponding probability is given.**

on improving the computation time by comparing with the traditional method.

An XGBoost-based prediction model is built and the model has achieved about 97% accuracy on test data. Technically, it is satisfactory. But we want to understand why a certain case of disturbance in power system is predicted to be unstable and why others are not. An electrical engineer would then be able to assess whether what the model learned makes intuitive sense and can be trusted. To achieve this, LIME is applied here. According to table 4, the LIME output for the six most important features within the two cases are shown in the Figure 12.

The artificial intelligence method used for transient stability prediction of power system is often a black box, for it does not go deep into the mechanism of power system. However, the method in this paper tries to explain the prediction results for system operators who are familiar with the mechanism of power system. The interpretable predictions is important to make humans trust and use machine learning effectively, if the explanations are faithful and intelligible. The Figure 12 shows the explanation for the prediction of the two cases. The first figure shows the case 1 (fault occurs at line 5-7), which is classified as instability with 99% probability. Here, the close relationship between the weight of features and the operating state of system in time domain is constructed. The feature $f_9$ is the difference of maximum and minimum generator rotor angle at the cutting time of the fault. The large difference between rotor angles will result in large weight for the value of $f_9$ to predict the case as ''instability''. The feature $f_5$ is the total system ''energy adjustments''. In general, the power system tends to be unstable while the value of $f_5$ is large. The feature $f_{13}$ is the total mechanical power before the fault incipient time and it indicates the load level which represents the general static stability level. All features have different weights and they are combined to predict transient stability accurately.

The process of interpreting individual predictions is illustrated in the Figure 12. It is much easier for the system operator to position the status and make a decision with the help of a model if intelligible explanations are provided. In this case, an explanation is about some import features with relative weights that either contribute to the prediction (in green) or are evidence against it (in red). System operators usually have prior knowledge about the power system, which they can use to accept or reject a prediction if they understand the reasoning behind it. A model predicts that the case is stable, and LIME highlights the features among other features that led to the prediction. The values of $f_9$, $f_5$, $f_{14}$ and $f_{18}$ are portrayed as contributing to the ''instability'', while the values of $f_{13}$ and $f_{10}$ are evidence against it. By this process, a system operator can make an informed decision about whether to trust the model's prediction.

## VII. CONCLUSION

This paper presents a transient stability prediction method for power system based on XGBoost. Firstly, the important features of generator operating state during transient process have been extracted by analyzing the dynamics of generator. Meanwhile, the redundant features are removed based on the correlation filtering and the model-based feature selection. Then, the relationship between the features and transient stability has been explained by the decision rules and feature importance scores. At last, the XGBoost model constructed is used to predict the transient stability based on the selected features under the specific operating cases. By the simulation results, the following conclusions can be drawn.

(1) By comparing with the traditional method, more features can be considered in the XGBoost-based method, which is attractive in complex power system, especially in the power system with new energy source penetrated.

(2) By this proposed method, the important features can be extracted among the massive features, which has the most direct correlation with transient stability.

(3) As the outstanding advantage for the XGBoost-based method, the fast computation makes it more possible to apply in the online prediction of a comprehensive power system.

So the further research for this method will be as follows.

(1) In order to make the model easier to understand, some more interpretative initial features can be constructed.

(2) Other hyper-parameter optimization approaches (e.g. random search) can be introduced in the XGBoost model to improve its adaptability in comprehensive power system.

(3) After the judgment for system stability, the deep learning methods can be introduced to combine with power system to locate the fault when the power system suffers fault.

## REFERENCES

[1] Z. Pannell, B. Ramachandran, and D. Snider, ''Machine learning approach to solving the transient stability assessment problem,'' in *Proc. IEEE Texas Power Energy Conf. (TPEC)*, College Station, TX, USA, Feb. 2018, pp. 1–6.

[2] J. Shu, W. Xue, and W. Zheng, "A parallel transient stability simulation for power systems," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1709–1717, Nov. 2005.

[3] P. Bhui and N. Senroy, "Real-time prediction and control of transient stability using transient energy function," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 923–934, Mar. 2017.

[4] E. Chiodo and D. Lauria, "Transient stability evaluation of multimachine power systems: A probabilistic approach based upon the extended equal area criterion," *IEE Proc.-Gener., Transmiss. Distrib.*, vol. 141, no. 6, pp. 545–553, Nov. 1994.

[5] J. D. E. Echeverría, C. J. C. Cepeda, and D. G. Colomé, "Real-time transient stability assessment of electric power systems using predictive-SIME based on machine learning," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf.-Latin Amer. (ISGT Latin America)*, Quito, Ecuador, Sep. 2017, pp. 1–6.

[6] D. Q. Zhou, U. D. Annakkage, and A. D. Rajapakse, "Online monitoring of voltage stability margin using an Artificial Neural Network," in *Proc. IEEE PES Gen. Meeting*, Minneapolis, MN, USA, Jul. 2010, p. 1.

[7] A. Karami, "Power system transient stability margin estimation using neural networks," *Int. J. Elect. Power Energy Syst.*, vol. 33, no. 4, pp. 983–991, 2011.

[8] M. Arefi and B. Chowdhury, "Post-fault transient stability status prediction using grey wolf and particle swarm optimization," in *Proc. SoutheastCon*, Charlotte, NC, USA, Mar./Apr. 2017, pp. 1–8.

[9] F. Gomez, A. Rajapakse, U. Annakkage, and I. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Detroit, MI, USA, Jul. 2011, p. 1.

[10] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, "Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs," *Int. J. Elect. Power Energy Syst.*, vol. 44, no. 1, pp. 318–325, 2013.

[11] P. M. Devie and S. Kalyani, "Transient stability prediction in multimachine system using data mining techniques," *Int. J. Eng. Res. Technol.*, vol. 2, no. 12, pp. 2748–2755, 2013.

[12] A. Y. Abdelaziz and M. A. El-Dessouki, "Transient stability assessment using decision trees and fuzzy logic techniques," *Int. J. Intell. Syst. Appl.*, vol. 5, no. 10, pp. 1–10, 2013.

[13] T. Amraee and S. Ranjbar, "Transient instability prediction using decision tree technique," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3028–3037, Aug. 2013.

[14] M. Rahmatian, Y. C. Chen, A. Palizban, A. Moshref, and W. G. Dunford, "Transient stability assessment via decision trees and multivariate adaptive regression splines," *Electr. Power Syst. Res.*, vol. 142, pp. 320–328, Jan. 320.

[15] C. Zhang, Y. Li, Z. Yu, and F. Tian, "Feature selection of power system transient stability assessment based on random forest and recursive feature elimination," in *Proc. IEEE PES Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, Xi'an, China, Oct. 2016, pp. 1264–1268.

[16] C. Zhang, Y. Li, Z. Yu, and F. Tian, "A weighted random forest approach to improve predictive performance for power system transient stability assessment," in *Proc. IEEE PES Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, Xi'an, China, Oct. 2016, pp. 1259–1263.

[17] V. Thirugnanasambandam and T. Jain, "AdaBoost classifiers for phasor measurements-based security assessment of power systems," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 8, pp. 1747–1755, Apr. 2018.

[18] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.

[19] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*, vol. 7. New York, NY, USA: McGraw-Hill, 1994.

[20] J. Lu, L. Lin, and H. Li, "Power system transient stability feature filter based on grey incidence clustering," in *Proc. 5th Int. Conf. Crit. Infrastruct. (CRIS)*, Beijing, China, Sep. 2010, pp. 1–5.

[21] Z. Pannell, B. Ramachandran, and D. Snider, "Machine learning approach to solving the transient stability assessment problem," in *Proc. IEEE Texas Power Energy Conf. (TPEC)*, College Station, TX, USA, Feb. 2018, pp. 1–6.

[22] I. B. Mustapha and F. Saeed, "Bioactive molecule prediction using extreme gradient boosting," *Molecules*, vol. 21, no. 8, p. 983, 2016.

[23] Y. Xia, C. Liu, Y. Li, and N. Liu, "A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring," *Expert Syst. Appl.*, vol. 78, pp. 225–241, Jul. 2017.

[24] M. Luckner, B. Topolski, and M. Mazurek, "Application of XGBoost algorithm in fingerprinting localisation task," in *Proc. IFIP Int. Conf. Comput. Inf. Syst. Ind. Manage.* Cham, Switzerland: Springer, 2017, pp. 661–671.

[25] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?': Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, 2016, pp. 1135–1144.

**MINGHUA CHEN** received the B.S. degree in automation from the Shandong University of Technology, Zibo, China, in 2016. He is currently pursuing the M.S. degree with the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His current research interests include power system security, stability analysis, and machine learning.

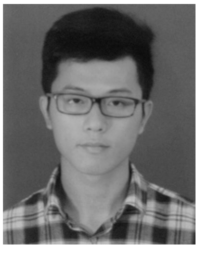
**QUNYING LIU** (M'13) received the B.Eng. and M.S. degrees and the Ph.D. degree in electrical engineering from Sichuan University, Chengdu, China, in 2000, 2005, and 2008, respectively. She is currently with the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu.

Her current research interests include power system security and stability analysis and advanced application of PMUs.

**SHUHENG CHEN** received the B.Eng. and M.Sc. degrees from the Harbin Institute of Technology, Harbin, China, in 1998 and 2000, respectively, and the Ph.D. degree from Sichuan University, Chengdu, China, in 2007, all in electrical engineering.

From 2013 to 2014, he was a Visiting Scholar with the Department of Energy Technology, Aalborg University, Aalborg, Denmark. Since 2008, he has been a Lecturer with the School of Energy Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His research interests include distribution system analysis, reconfiguration, and reactive power optimization.



**YICEN LIU** received the B.S. degree in automatic from the University of Electronic Science and Technology of China, Chengdu, China, in 2017, where he is currently pursuing the M.S. degree in control engineering.

His research interests include the transient analysis of power systems with wind power, and power system modeling with energy function.

**CHANG-HUA ZHANG** (M'09) was born in Hubei, China. He received the Diploma degree in vehicle engineering from the Wuhan University of Science and Technology, Wuhan, China, in 1993, the M.Sc. degree in vehicle engineering from the Wuhan University of Technology, Wuhan, in 2004, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, in 2007. From 2011 to 2012, he was a Visiting Scholar with Loughborough University, U.K.

He is currently an Associate Professor with the School of Mechanical and Electrical Engineering, University of Electronics Science and Technology of China. His current research and academic interests include smart grid, converter control, and microgrids analysis theory.

**RUIHUA LIU** received the B.Eng. degree from Shanxi University, Taiyuan, China, in 2006, and the M.S. degree from Sichuan University, Chengdu, China, in 2009. She is currently with the Skills Training Center, State Grid Sichuan Electric Power Company.

Her main research interests include power system automation and power market.